

DOI: 10. 20040/j. cnki. 1000-7709. 2023. 20220494

# 基于 IEWT-FE-BO-LSTM 模型的超短期风功率预测

陆秋贤<sup>1</sup>, 马 刚<sup>1</sup>, 涂孟夫<sup>2</sup>

(1. 南京师范大学电气与自动化工程学院, 江苏 南京 210023; 2. 南瑞集团有限公司, 江苏 南京 211106)

**摘要:** 为提高超短期风功率预测精度, 提出一种基于 IEWT-FE-BO-LSTM 的组合风功率预测模型, 首先利用改进经验小波分解(IEWT)对历史风功率数据进行分解; 然后引入模糊熵(FE)算法对各分解子模态进行复杂度计算重组子模态; 再对各个重组分量分别建立基于长短时神经网络(LSTM)的预测模型, 利用贝叶斯优化算法(BO)进行超参数组合, 解决人为调参导致训练结果不佳的问题; 最后通过历史风电场数据进行算例分析。结果表明, IEWT-FE-BO-LSTM 模型对超短期风功率有较高的预测精度和预测效率。

**关键词:** 超短期风功率预测; 改进经验小波分解; 模糊熵; 贝叶斯优化算法

**中图分类号:** TM614

**文献标志码:** A

**文章编号:** 1000-7709(2023)01-0217-04

## 1 引言

“十三五”以来, 全国累计并网风电装机容量逐年提高<sup>[1]</sup>, 风功率预测可有效地制定电力调度计划, 从而在一定程度上提高风电的市场竞争力, 因此提高风功率预测的精度更具有现实意义<sup>[2]</sup>。风功率序列具有不确定性、波动性大的特点, YAN JIANG 等<sup>[3]</sup>利用波形分解与机器学习算法相结合的组合预测方法将风功率序列进行分解从而可有效提高预测精度<sup>[4]</sup>, 克服风功率的波动性和非平稳特性。目前主要的分解方法有经验模态分解(EMD)、完全自适应噪声集合经验模态分解(CEEMDAN)、经验小波分解(EWT)等, EMD 已被验证在分解时会产生模态混叠现象; CEEMDAN 将含噪声较多的分量直接舍弃, 易造成信息缺失; EWT 方法处理频谱较复杂的信号容易因边界划分不当, 从而频谱泄漏和噪声污染带来的局部无效峰值导致“过切分”现象, 对于模态数难以把控从而影响组合预测模型的效率。为此, 本文提出一种基于 IEWT-FE-BO-LSTM 的组合风功率预测模型, 即采用改进经验小波方法, 在傅里叶谱上根据试验数据集制定划分规则, 确定各带通滤波器的边界, 剔除噪声分量并利用模

糊熵(FE)重构子模态从而提高整体预测效率, 选用 LSTM<sup>[5]</sup>作为基础训练模型, 但其人为调试训练参数缺乏可靠性且计算效率低, 引入 BO 算法, 其在未知黑箱函数的情况下, 通过猜测黑箱函数求一个可接受的最大值, 自适应调节模型的参数以进行优化, 并通过历史风电场数据进行仿真验证。结果表明, 该组合预测模型解决了传统分解易发生模态混叠的同时, 也解决了 EWT 方法容易出现的频谱“过分切”现象, 提高了模型的预测效率及精度。

## 2 研究方法

### 2.1 IEWT 方法

虽然 EWT 方法<sup>[6]</sup>克服了 EMD、EEMD 方法由于信号时频尺度不连续引发的模态混叠问题及过包络和欠包络的问题, 但在处理频谱较复杂的信号容易出现频谱“过切分”现象, 所以本文在 EWT 原理基础上进行改进, 利用局部最小最大值方法预先对频谱包络进行峰值检测从而获取频谱包络线, 将各谷值之间的峰值区域进行预划分, 并设定如下规则对分割边界进行调整的同时消除噪声干扰: ①跨度大于总长度的  $\frac{1}{10n}$  ( $n$  为划分区

收稿日期: 2022-03-16, 修回日期: 2022-04-22

基金项目: 2020 年江苏省重点研发计划(BE2020081-4)

作者简介: 陆秋贤(1998-), 男, 硕士研究生, 研究方向为新能源接入下机组组合方法, E-mail: 1826375641@qq.com

通讯作者: 马刚(1984-), 男, 博士、副教授, 研究方向为新能源发电及入网技术、综合能源系统、电力系统分析及故障诊断相关技术, E-mail: nnumg@njnu.edu.cn

间数量,根据数据复杂度预先设定)或频带中的最大值小于前一频带平均值的频带为噪声分量。②噪声分量不多于子模态数量的 1/4。③子模态数量不多于传统 EMD 方法所得分量数的 2 倍。改进过后的 IEWT 方法更具有灵活性和鲁棒性。

### 2.2 FE 算法

虽然 EWT 和 IEWT 方法较 EMD 方法克服了模态混叠及过包络和欠包络的问题,但是传统 EWT 方法应用于分解风功率数据序列这样不确定性、波动性大的数据序列将得到更多的子模态,从而增大整体预测复杂度。为了提高模型预测效率,从考量各子模态相关性切入,FE 对传统 EWT 方法分解所得子模态分别计算复杂度,考量其相关性从而对各子模态进行重组。FE 衡量新模式产生的概率大小,测度值越大,新模式产生的概率越大,即序列复杂度越大<sup>[7]</sup>。对给定  $N$  维时间序列  $X$  进行相空间重构,得到时间序列  $Y$ :

$$Y(i) = [x(i), x(i+1), \dots, x(i+m-1)] - x_0(i) \quad (1)$$

式中,  $m$  为嵌入维。

定义时间序列  $Y(i)$ 、 $Y(j)$  之间距离为:

$$d_{i,j}^m = \max_{k \in (0, m-1)} | (x(i+k) - x_0(i)) - (x(j+k) - x_0(j)) | \quad (2)$$

式中,  $k$  为  $0 \sim m-1$  的任意数。

引入模糊隶属函数,利用模糊函数计算  $Y(i)$ 、 $Y(j)$  之间相似度  $D_{i,j}^m$ :

$$D_{i,j}^m = \exp[-(d_{i,j}^m)^n / r] \quad (3)$$

式中,  $r$  为相似容限度;  $i, j = 1, 2, \dots, N-m+1$ , 且  $i \neq j$ 。

定义:

$$\psi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \left( \frac{1}{N-m} \sum_{j=1, j \neq i}^{N-m+1} D_{i,j}^m \right) \quad (4)$$

式中,  $N$  为  $N$  维数列。

原时间序列 FE 为:

$$\text{Fuzzy}E_n(m, r) = \lim_{N \rightarrow \infty} [\ln \psi^m(r) - \ln \psi^{m+1}(r)] \quad (5)$$

有限时间序列 FE 估计为:

$$\text{Fuzzy}E_n(m, r, N) = \ln \psi^m(r) - \ln \psi^{m+1}(r) \quad (6)$$

### 2.3 BO-LSTM 时间序列预测

#### 2.3.1 贝叶斯优化算法

贝叶斯优化有两个核心过程,即先验函数(PF)与采集函数(AC);AC 用来完成探测与开发的平衡,主要包括 EI(Expected Improvement)、PI(Probability of Improvement)、UCB(Upper

Confidence Bound)几种函数。PI 原理为找到未知点的函数值比  $f(x^+)$  大的概率,取这些点中概率最大的点,但不考虑差值。EI 函数求得未知点函数值比  $f(x^+)$  大的数学期望,解决了 PI 未考虑未知点与最大点差值的问题<sup>[8]</sup>。

#### 2.3.2 基于 BO 的 LSTM 模型

选用 BO 算法对 LSTM 超参数进行优化,BO 算法的核心思想在于利用 PF 逼近未知目标函数的后验分布,然后根据分布选择下一个采样的超参数组合。为尽量接近真实目标函数,BO 采用代理模型(主要使用高斯过程回归)。高斯过程回归算法定义了 PF,可用来合并关于目标函数的先验信息(平滑度等)。本文方法流程见图 1。

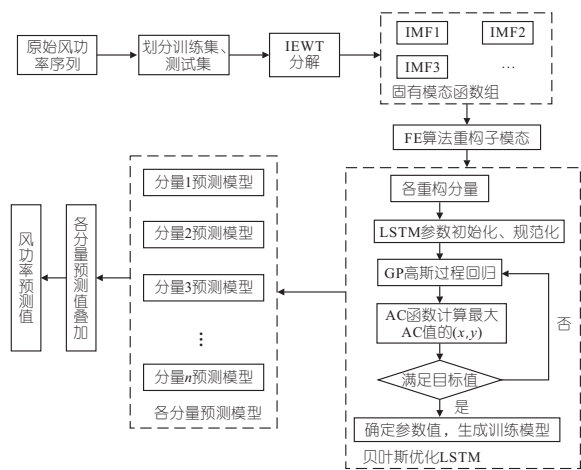


图 1 IEWT-FE-BO-LSTM 预测流程图

Fig. 1 IEWT-FE-BO-LSTM prediction process diagram

### 2.4 模型评价指标

(1) 均方根误差 ( $R_{RMSE}$ ):

$$R_{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (7)$$

式中,  $m$  为预测量数量;  $y_i$  为实际值;  $\hat{y}_i$  为预测值。

(2) 平均绝对百分比误差 ( $M_{MAPE}$ ):

$$M_{MAPE} = \left( \frac{1}{m} \sum_{i=1}^m \left| \frac{y_i - \hat{y}_i}{y_i} \right| \right) \times 100\% \quad (8)$$

(3) 均方百分比误差 ( $M_{MSPE}$ ):

$$M_{MSPE} = \sqrt{\frac{1}{m} \sum_{i=1}^m \left( \frac{y_i - \hat{y}_i}{y_i} \right)^2} \quad (9)$$

## 3 算例分析

#### 3.1 样本生成与预处理

本文所有算例仿真均在 Intel Core i7-9700K CPU @ 3.60G Hz、16GB RAM、RTX2080 的硬件环境下进行测试仿真。仿真软件使用

Matlab2019a 对神经网络进行搭建。采用江苏省某风电场所提供的 2016 年 1~2 月风功率出力实际数据作为数据集来划分训练集和测试集,其采样周期为 15 min。为体现本文预测方法的应用广泛性,在数据集中随机选取一组连续数据,选取 400 个风功率数据点作为样本,其中前 300 个数据点作为训练集,后 100 个作为测试集以验证模型仿真的预测结果,功率曲线见图 2。

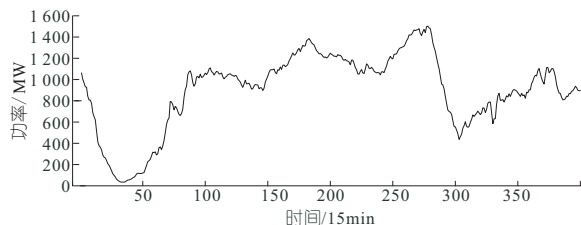


图 2 风功率曲线图

Fig. 2 Wind power curve

### 3.2 预处理模型

#### 3.2.1 IMF 分量最大数量预设

为验证 EWT 分解的 IMF 分量个数的合理性,利用 EWT-LSTM 模型预先对分量数设置为 5、6、7、8、9 分别建模仿真对照,验证结果见表 1。由表 1 可知,5 个 IMF 分量无法有效利用 EWT 方法进行最优分解,而随着分量数的增加,其在最终结果的表现上差距不大,但设置为 8 个分量明显最优,所以本文 EWT 方法的 IMF 分量最大数量预设 8。

表 1 IMF 数量设置合理性对比

Tab. 1 Comparison of rationality of IMF quantity setting

分量数	$R_{RMSE}$	分量数	$R_{RMSE}$
5	20.437 9	8	13.413 4
6	14.597 0	9	14.247 1
7	14.597 1		

#### 3.2.2 IEWT 分解分析

利用本文 IEWT 方法对风功率序列进行预分解确定其频谱划分边界,划分结果见图 3,对原始风功率序列进行分解,得到 7 个 IMF 分量和一个余量,按照其分解所得模态频率大小可分为高、中、低频。同时为验证方法的优越性,使用文献 [9] 中提出的 CEEMDAN 及 EMD 方法对原始风

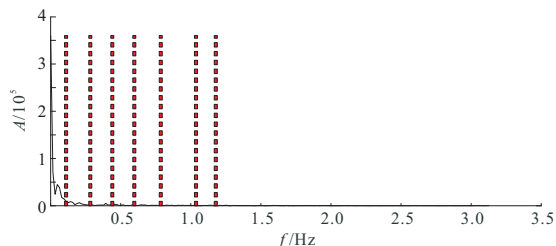


图 3 IEWT 频谱边界划分

Fig. 3 Division of IEWT spectrum

功率序列进行分解,分别得到 8 个 IMF 分量、1 个余量和 5 个 IMF 分量、1 个余量。IEWT 法子模块数明显少于 CEEMDAN 方法,即需建立子预测模型的数量更少,预测效率更优。对比可见,IEWT 方法减小了模态混叠现象,有效分离了高频振荡及低频基础信号。IEWT 的高频信号出现在了 IMF3~IMF8,而高频信号的幅值波动明显小于 EMD 分解的高频信号 IMF1、IMF2,可见 IEWT 分解可将高频大幅信号有效分解为高频小幅信号,从而降低了各 IMF 分量预测的难度。

#### 3.2.3 FE 子模块重构

本文设置 FE 参数  $n$  为 2,参数  $r$  为 0.15,其重构结果见表 2。由表 2 可知,将 8 个原始 EWT 子模态重构为 6 个新分量,降低了子预测模型的数量,可提高其总预测效率。

表 2 子模块重构

Tab. 2 IMF reconfiguration

重构分量	初始模态分量	重构分量	初始模态分量
分量 1	IMF6、IMF7、IMF8	分量 4	IMF3
分量 2	IMF1	分量 5	IMF4
分量 3	IMF2	分量 6	IMF5

### 3.3 IEWT-FE-BO-LSTM 组合超短期风功率预测

对 IEWT 算法所分解的各个 IMF 分量经过 FE 重构分量后分别建立 LSTM 时序预测模型,LSTM 设置提前 2 步预测,最大迭代次数设置为 300,根据测试各优化器选用下误差对比后选用 Adam 算法优化器,其模型的收敛速度最快,预测误差指标表现最佳,具体各优化器误差指标表现结果见表 3,最后利用 BO 算法对 LSTM 进行优化以自适应进行参数选择。

表 3 各 LSTM 优化器选择下预测误差指标对比

Tab. 3 Comparison of prediction error indicators under the selection of each LSTM optimizer

算法优化器	$R_{RMSE}/\%$	$M_{MAPE}/\%$	$M_{MSPE}/\%$
Rmsprop	12.183 0	1.662 9	1.945 6
Sgdm	11.883 9	1.225 5	1.579 1
Adam	11.378 0	1.130 5	1.458 6

为验证本文所提出模型的有效性,分别使用文献 [10] 中所提的传统预测模型 Elman、LSTM、EMD-LSTM、文献 [9] 中 CEEMDAN-LSTM、EWT-LSTM 及递增加入本文建模方法的模型预测风功率序列,并与本文所提模型进行对比,最终预测结果见表 4。

由表 4 可知,与单一 LSTM 模型预测相比,CEEMDAN-LSTM 组合模型预测的误差指标表现良好,明显降低了误差;而利用 EWT 波形分解,将高频大幅信号有效分解为高频小幅信号,从而降低了各 IMF 分量预测的难度,提高了预测精

表 4 各模型预测误差指标对比

Tab.4 Comparison of prediction error indexes of each model

模型	$R_{RMSE}/\%$	$M_{MAPE}/\%$	$M_{MSPE}/\%$
Elman	64.407 3	6.051 8	8.544 8
LSTM	37.719 2	3.658 3	1.133 4
BO-LSTM	32.366 6	3.819 5	0.889 1
EMD-LSTM	21.690 7	2.240 4	2.900 3
CEEMDAN-LSTM	19.555 7	1.802 3	2.630 8
EWT-LSTM	15.789 4	1.541 2	2.124 8
IEWT-LSTM	13.413 4	1.381 8	1.963 8
IEWT-BO-LSTM	11.378 0	1.130 5	1.458 6
IEWT-FE-BO-LSTM	11.569 4	1.216 8	1.676 2

度,而 IEWT 方法与传统 EWT 方法相比,有效根据数据结构控制了频谱划分,在结果上预测精度明显更高,验证了 IEWT 方法的有效性,与 EMD-LSTM 组合预测模型相比,其  $R_{RMSE}$ 、 $M_{MAPE}$ 、 $M_{MSPE}$  分别降低了 38.16%、38.22%、32.29%,而文献[9]方法虽在 EMD-LSTM 的基础上有所改善,但精确程度离本文方法仍有差距,验证了 IEWT 方法对于风功率数据这种不确定性波动大的时间序列数据预测的有效性。经过 BO 算法优化的模型预测误差明显降低,在 IEWT-LSTM 的基础上,IEWT-BO-LSTM 的  $R_{RMSE}$ 、 $M_{MAPE}$ 、 $M_{MSPE}$  分别降低了 15.17%、18.19%、25.73%,对比其他预测方法,其预测精度得到大幅改善,验证了本文方法的有效性。在重构 IMF 分量后所搭建的 IEWT-FE-BO-LSTM 模型与重构分量前的预测模型相比,由于经过了模态重组,降低了分量数,减少了 2 个分量训练模型,从而提高了预测效率,且在预测精度上也仅有 1.6821% 的误差,在保证预测精度的同时提高了模型的总体预测效率。最终风功率预测曲线见图 4。

## 4 结论

本文提出的基于 IEWT-FE-BO-LSTM 的组合风功率预测模型可有效解决传统分解方法的模态混叠现象及频谱“过分切”现象,控制了子模

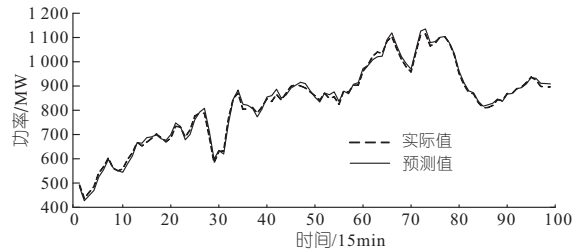


图 4 超短期风功率预测曲线图

Fig.4 Ultra short term wind power prediction curve

态数量,在保证预测精度的同时提高了模型的总体预测效率。

## 参考文献:

- [1] 祁鑫,韩红卫,涂孟夫,等. 大规模新能源接入下基于概率分析的调度计划及安全校核关键技术研究总技术报告[R]. 银川:国网宁夏电力公司, 2017.
- [2] 赵征,汪向硕. 基于 CEEMD 和改进时间序列模型的超短期风功率多步预测[J]. 太阳能学报, 2020, 41(7):352-358.
- [3] YAN JIANG, GUOQING HUANG. Short-term wind speed prediction: Hybrid of ensemble empirical mode decomposition, feature selection and error correction[J]. Energy conversion & management, 2017, 144: 340-350.
- [4] 张群,唐振浩,王恭,等. 基于长短时记忆网络的超短期风功率预测模型[J]. 太阳能学报, 2021, 42(10): 275-281.
- [5] 李卓,叶林,戴斌华,等. 基于 IDSCNN-AM-LSTM 组合神经网络超短期风电功率预测方法[J]. 高压技术, 2022, 48(6):2117-2127.
- [6] 李红辉,周建中,张勇传,等. 基于 EWT 和修正 Morris 法的抽水蓄能电站主进水阀自激振动研究[J]. 振动与冲击, 2019, 38(19):52-57,76.
- [7] 杨维奇. 基于脑机接口的脑电分类算法研究与实现[D]. 四川:电子科技大学, 2019.
- [8] 陈娜. 基于机器学习的智能电网短期电价预测算法研究[D]. 上海:东华大学, 2021.
- [9] 张景杨,王维庆,王海云,等. 基于 CEEMDAN-SAFA-LSSVM 的短期风功率预测[J]. 计算机仿真, 2021, 38(8):134-138.
- [10] 杨楠,叶迪,周峥,等. 基于 NACEMD-Elman 神经网络的风功率组合预测[J]. 水电能源科学, 2018, 36(9):209-211,171.

## Ultra-short-term Wind Power Prediction Based on IEWT-FE-BO-LSTM Model

LU Qiu-xian<sup>1</sup>, MA Gang<sup>1</sup>, TU Meng-fu<sup>2</sup>

(1. School of Electrical and Automation Engineering, Nanjing Normal University, Nanjing 210023, China; 2. NARI Group Co. Ltd., Nanjing 211106, China)

**Abstract:** In order to improve the prediction accuracy of ultra short-term wind power, a combined wind power prediction model based on IEWT-FE-BO-LSTM was proposed. Firstly, an improved empirical wavelet decomposition (IEWT) was used to decompose the historical wind power data. The Fuzzy Entropy (FE) algorithm was introduced to calculate the complexity of each decomposed submodel and reconstruct the submodel. For each reconstructed component, a prediction model based on long short term neural network (LSTM) was established. Bayesian optimization algorithm (BO) was used for hyper parameters to solve the problem of poor training results caused by manual parameter adjustment. The example analysis based on historical wind farm data shows that the IEWT-FE-BO-LSTM model has high prediction accuracy and efficiency for ultra short-term wind power.

**Key words:** ultra short-term wind power prediction; improved empirical wavelet transform; fuzzy entropy; Bayesian optimization algorithm